6

Conclusions/Recommendations

6.1 Historical Assessment Comparisons

Comparisons presented in this document are of three types:

- RBP results with different types of historical data: Hester-Dendy multiplate samplers (Ohio) and the traveling kick net (New York).
- RBP sampling with variation of taxonomic level (New York).
- RBP sampling with variation of subsample size (New York).

A comparison of results suggested a reasonably good fit between Ohio EPA findings and those of the present study. Subtle discrepancies between the data sets are most likely a result of the lack of regional calibration for the RBP analysis technique; that is, there is not a complete understanding of which benthic metrics are most appropriate for the upper Midwest when using kick nets. This might have weakened the interpretive power of the approach. Also, there is likely some effect of the different sampling methodologies (Hester-Dendy multiplate samplers and square-meter kick nets) on the assessments. It is difficult to determine if these more subtle differences are due to differences in methods or changes in biological condition over time. Bioassessment, as exemplified by the Ohio EPA ICI (for macroinvertebrates) and IBI (for fish) and the EPA RBP (for macroinvertebrates), is a valid and technically sound tool for evaluating impaired waters, particularly when calibrated on a regional level as is done for the ICI and IBI. This validation is supported by similar assessments being arrived at by approaches differing in detail (this study).

For the New York portion of the study, all assessments compared favorably with those most recently performed by the DEC (Bode et al. 1993). In 1990 sampling, Canastota Creek was found to be "moderately impacted" at a single station downstream of the town. At three stations along its length, we assessed it as "slightly to moderately impaired" and "moderately impaired."

A downstream station on Onondaga Creek was assessed as "severely impacted" in both 1989 and 1990 sampling efforts (Bode et al. 1993). Our assessments showed this creek to be "moderately impaired" in upstream reaches and "severely impaired" near the same station assessed by DEC. Harbor Brook was assessed similarly between DEC in 1989 (Bode et al. 1993) and here as "moderately" to "severely impaired."

Traditional comparisons of biological assessment methods occur through side-by-side sampling and analysis. These temporally separate data have provided some useful insights into the process of bioassessment comparisons. As mentioned above, differences in results might arise directly from sampling biases inherent in the sampling gear. This might be a problem when attempting to directly compare data from separate bioassessment samples (e.g., the number of species, the calculated value of an individual metric or the number of individual organisms collected). The problem of sampling error (bias) is reduced if comparisons are made at the level of the overall assessment score rather than individual metrics.

6.2 Statistical Comparisons

Comparisons were made between RBPII (family-level identifications) and RBPIII (lowest-practical-level identifications, usually genus/species), as well as subsample size (100-organism versus 300-organism). As long as the reference conditions are treated in the same manner as test station data (taxonomic and subsampling levels), comparisons between assessment results are valid. We found that although there might have been some differences in specific metric performance (i.e., metric values) with different treatments, those values relative to reference conditions varied little. Further, there was perfect agreement among total bioassessment scores between the treatments.

6.2.1 Taxonomic Level Conclusions

When addressing the question of appropriate taxonomic level, different concerns do arise. Although similar site rankings based on condition might be found with different levels, there can be difficulty in interpretation of potential causative factors when using more gross-level identifications. This is especially true when dealing with metrics dependent on how individual species adapt to the environment rather than how they relate to other species. These metrics include the HBI and those related to functional feeding groups (scrapers, filterer collectors, shredders). The tolerance values on which the HBI is based are usually assigned to species (or genus) level and might not be available for family. Likewise, functional feeding group designations become more uncertain as they are assigned to more general (or higher) taxonomic levels. It is recommended that, in general, taxonomy be performed to the lowest practical level that will suit the objectives of the study, which will usually be the genus or species level for biological assessments beyond the screening level. The decision on taxonomic level might also be refined with regional calibration of bioassessment techniques.

6.2.2 Subsample Size Conclusions

The argument can be made that a lower number of organisms does not allow a reasonable estimate of biological diversity. However, as was shown with the taxa richness metric, as higher numbers of organisms are included in a sample, the higher the number of detected taxa will be. This is due to an increase in the probability of rare taxa being included within a larger subsample. In essence, rare taxa have little influence on biological assessments using a multimetric approach because even if rare taxa are collected, their contribution to a multimetric index is minimal. Conversely, if one's goal is to describe biological diversity at a site, even an analysis of the total sample (versus a subsample) is likely inadequate. It is possible to collect continuously larger samples from a broader diversity of microhabitats within a site and continue to get additional taxa. The critical factors are to have consistency in sampling effort and a properly randomized subsampling procedure. As with other sample treatments, subsampling is appropriate as long as samples from reference sites are treated in the same way; subsamples less than 100-organisms are not recommended. The recommendation is to base benthic macroinvertebrate biological assessments on 100-organism subsample when using RBPs in New York.

6.3 Usefulness of RBPs in Assessing CSO Biotic Effects

Attributing cause and effect to the specific CSO activity is complicated by other related problems associated with urbanization, e.g., habitat alteration and industrial discharges. However, the bioassessment procedures, with its integration of total scores, individual metrics (which are based on known ecology of the benthic community) and habitat description, provide reasonable technical support for identifying potential sources of biological impairment. An impairment due to CSO outfalls was noted in biological data collected by both Ohio EPA and the present study for a 15- to 20-mile reach of the Scioto River, a 4-mile reach of the Sandusky River, and a 10-mile reach of the Little Cuyahoga River. In the cases of the Scioto and Little Cuyahoga Rivers, upstream stations also located in urbanized areas had relatively healthy biological communities and were effective for comparisons of biological data. The unimpaired middle station of the Little Cuyahoga River exhibited recovery of the biota since the correction of upstream CSOs. The assessments were performed prior to our gaining information concerning the outfalls.

For the New York study, severe habitat degradation and alterations were evident at all Onondaga Creek sites and at the two downstream sites on Harbor Brook. There were many instances of major habitat differences between stations on the same stream or between a station and its regional reference site. However, even with these differences, impairment due to stressors commonly produced by CSOs was seen at the middle and lower stations on Canastota Creek and Onondaga Creek.

Results indicated that CSO outfalls had an adverse impact on the downstream macroinvertebrate assemblages. Impairment of the benthic biota, in both the Ohio and New York studies, was manifested by the metrics (1) taxa richness, (2) scraper/scraper + filterer collector, (3) EPT/EPT + Chironomidae, (4) percent contribution of dominant taxon, (5) Hydropsychidae/total Trichoptera, (6) Pinkham-Pearson Community Similarity Index, (7) QSI-taxa, and (8) DIC-5.

The bioassessments were instrumental in identifying impaired reaches of each river at periods that reflected residual and cumulative effects of CSO outfalls that were not necessarily actively discharging. Sampling was performed during normal flow conditions (i.e., not during the wet or dry season) although several of the Ohio sampling locations were being affected by increased flow levels. Results illustrate the utility of biological data for capturing the effects of intermittent discharge events without sample collection during stormflows.

The use of multiple metrics aids in achieving more accurate assessments than single-parameter assessments. This was seen in the case of HB1 and HB2, which had nearly identical metric values for taxa richness but very different overall biological assessments (HB1-moderate, HB2-severe). The multimetric approach uses the total assessment score for comparison to the reference in determining the biological integrity at a site and uses individual metrics for interpreting the assessment and gaining insight as to cause-and-effect relationships. The associated habitat assessment enabled a characterization of the physical habitat alteration, strengthening the ability to identify additional potential sources of impairment. For example, the nonimpaired biological condition assessment in the presence of degraded habitat on the Sandusky River (SA2) is a likely indication of some form of nutrient enrichment since, as discussed earlier (Section 3.2.3.4), the initial phases of nutrient enrichment cause an increase in the biota. If the nutrient enrichment is mild to moderate, the biological community balances between the effects of enhanced biota and the next phase of enrichment, oxygen depletion. In such instances, the biology would continue to score higher than the surrounding habitat would be expected to support.

6.4 The Place of Bioassessment In Watershed Protection

Another potential application for bioassessments is within the total maximum daily load (TMDL) process, which is one of the essential tools of the watershed protection approach. The watershed protection approach attempts to evaluate watersheds on a holistic, rather than piecemeal, basis. A TMDL is defined by USEPA guidance and regulations as being equivalent to the loading capacity of a waterbody and the sum of the individual wasteload allocations (WLAs) for point sources, load allocations (LAs) for nonpoint sources, and natural background sources, and a margin of safety to account for uncertainties about the relationship among stressors, controls, and the quality of the receiving water (USEPA 1994b).

TMDLs are required when states determine that technology-based controls will not result in a waterbody's meeting water quality standards, including its designated uses. The TMDL process can provide sufficient and necessary information for making decisions on the implementation of appropriate pollution reduction tools such as best management practices (BMPs), ecological restoration, or engineered active or passive treatment technologies (USEPA 1994c).

Although TMDLs until now have been primarily chemical-specific, biological assessment shows promise as a tool for going beyond chemical water quality to biological endpoints and the aquatic life uses of the waterbody. Biological assessments provide a direct evaluation of ecosystem condition by integrating physical habitat quality with biological condition. The evaluation is accomplished by comparison to empirically-defined, regionalized expectations of biological conditions (reference conditions). As was demonstrated in these case studies, bioassessments can often detect the biological impact of CSOs and other intermittent discharges in urbanized watersheds affected by multiple stressors. Because CSOs contribute to the pollution load entering a waterbody, they must be considered in TMDL development. Biological assessment used in the TMDL process can help:

- Identify waters that are ecologically impaired and might be in nonattainment of chemical water quality standards; this would help in the siting and installation of appropriate controls.
- Prioritize and target ecologically impaired waters.
- Aid in the development and implementation of TMDLs for nonchemical stressors within a watershed.
- Assess the effectiveness of installed pollution control tools in protecting aquatic resources.
- Where the metrics for a region have been sufficiently refined, the diagnostic capabilities of some metrics might allow some conclusions to be drawn with regard to specific causes of biological impairment in a waterbody.

Other current USEPA programs that can benefit from the use of biological assessments include 1994 CSO Control Policy (section 1.1) , stormwater and wet-weather monitoring, 305(b) reporting, and biological criteria. Many states have incorporated biological assessments into their 305(b) reports, and many are currently developing biological criteria for waterbodies in their ecoregions. As illustrated in this report, biological assessments are useful for determining impairments from episodic events such as those accompanied by wet weather and stormwater without the necessity of sampling during the actual event.